



Yu, G. (2012, Nov 19). Relevance of EFL standardized tests to modern EFL classrooms and to teaching and learning environments. Oxford University Press.

Publisher's PDF, also known as Version of record

[Link to publication record in Explore Bristol Research](#)
PDF-document

Republished by the Journal of English Teachers Association Switzerland (2016) with OUP's permission

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Relevance of EFL standardized tests to modern EFL classrooms and to teaching and learning environments

Guoxing Yu,
Graduate School of Education
University of Bristol

(Guoxing.YU@bristol.ac.uk)

<http://www.bris.ac.uk/education/people/guoxing-yu/about.html>

Abstract

The “unhappy” relationships between standardized tests and classroom teaching and learning, aired by critics of standardized tests and advocates of classroom-based assessment from educational, political and social perspectives, can often be found in the mass media. However, standardized tests and classroom teaching and learning are not really irreconcilable. In many cases, I, as a language testing professional, will argue that standardized tests and classroom teaching and learning are, in fact, inseparable. Their connection can be compared to a marriage – it is never a particularly easy relationship, but they share a common interest in improving the processes and products of education and compromises often have to be made on both sides if this is to be achieved.

Arriving at a healthy, happy harmonious relationship between standardized tests, teaching and learning, requires real effort on the part of both the assessment and teaching communities to enhance better understanding and mutual dialogue. In this short paper, I suggest two areas where this kind of effort is needed:

- (1) test design and
- (2) integrating standardized performance test tasks into classroom instruction.

I will also argue that we can foster a healthier and happier relationship by working together to improve both the assessment literacy of language teachers and the classroom experiences of test designers as part of their respective professional development.

1. “Unhappy” relationships between standardized tests, classroom teaching and learning

Traditionally, standardized tests have been used for a multitude of educational decisions by administrators, teachers, students, parents and many other stakeholders (Rea-Dickins, 1997). For instance, test results are used to place students into different courses and programmes, to diagnose their learning needs, to determine whether learning goals have been met, to evaluate teaching methods and their effectiveness, and to provide self-evaluation information for teachers and learners at various stages. However, the relationship between standardized tests and classroom teaching and learning has been described as “unhappy” and even “irreconcilable” by many of those who criticise standardized tests and advocate classroom-based assessment.

Criticisms of this “unhappy” relationship have centred on the educational, political and social values associated with standardized tests, and the effect that they can exert at individual, classroom and societal levels. A student’s identity and (self-) identification can be affected by the tests that s/he takes, or the test results that s/he achieves and embodies. The test results may affect how a student thinks and behaves in a community of learning (Rea-Dickins, Kiely, & Yu, 2007). Teachers may be under pressure to try to squeeze in (or ironically take out) as much as possible in order to cover the material required for the test at the end of the course. When standardized tests are used for accountability purposes, students can get piegholed (e.g., as successes or failures) in rather a crude manner.

Academics have been researching the ways in which language tests may affect classroom teaching and learning. There are a substantial number of such studies, concerning what is often labelled as the “washback” or “impact” of language testing (e.g., Alderson & Wall, 1993; Bailey, 1996; Cheng, 1997; Cheng, Watanabe, & Curtis, 2004; Green, 2007; Hamp-Lyons, 1997; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Wall & Alderson, 1993, just to name a few), or the ethicality and effectiveness of intensive test preparation (Alderson & Hamp-Lyons, 1996; Green, 2006; Yu, in press), or language testing as a social practice (McNamara, 1998, 2001). These studies try to unpack (and perhaps sometimes accidentally aggravate) such “unhappy” relationships. However, a key question seems to be missing – does the relationship really have to be unhappy and irreconcilable? Can shared goals and common interests between testing, teaching and learning help to bring the communities together?

2. A house divided against itself cannot stand

Abraham Lincoln’s famous *House Divided Speech* (16 June, 1858) – “A house divided against itself cannot stand” – holds the wisdom to address the currently unhappy relationships between standardized tests, classroom teaching and learning. Perhaps testing and teaching will never enjoy a “happy ever after” fairy tale romance, but tests are and will remain a big part of what we do as teachers. We have to learn to live with them. Following the analogy of a marriage, guidance or counselling can help the partners to focus on shared goals to arrive at a reconciliation based on mutual understanding. The first step is for the assessment and teaching communities to really “talk” to each other like family members and to explore each other’s point of view.

2.1 Managing change in language tests by design

EFL teachers sometimes find that the standardized tests that they are asked to administer do not really reflect or connect to what is happening in their classrooms. They do not reflect current understanding of communicative language proficiency, language use in real life, or its teaching and learning in classroom contexts. If there is a mismatch between testing, teaching and learning it often seems that the standardized tests simply take over in the classroom. As a consequence, teachers may become disillusioned because test results may not accurately reflect what they know about their students’ language ability based on classroom performance. Teachers lose confidence in standardized tests, the use of the test results and the whole assessment system. Similarly, students may become de-motivated and less engaged with the assessment tasks as well as daily classroom learning activities.

The question to both assessment and teaching communities would be this: how can we better harness the power of standardized language tests to serve the multiple purposes of teaching, learning and assessment? A united house is not only desirable, but also achievable. Both tests and classroom teaching and learning activities should reflect current understandings of communicative language ability and language learning; and reciprocally tests, by design, can act as an important lever for change (Kellaghan & Greaney, 2001). Tests, like any other form of educational assessment, should involve activities that mirror and so encourage communicative language teaching and learning. In this way, perhaps any conflict between teaching for communication and ‘teaching to the test’ would become largely irrelevant. Taking the test could itself be viewed as a valuable part of the learning process. Teaching to the test would no longer be considered as unethical, because giving or taking standardized tests would become equivalent to teaching or learning; radically different from the current problems of test-driven teaching. In order to achieve this ideal situation, language testers need to learn from teachers about what makes an effective communicative task and design the assessment tasks in such a way that they should possess the quality of being readily useable as teaching and learning tasks in classroom contexts.

2.2 Integrating performance tasks in classroom teaching and learning

In the current practices of language assessment, it is speaking and writing assessment tasks that are arguably the most adaptable and useable for communicative language teaching and learning in classroom contexts. Performance assessments “attempt to emulate the context or conditions in which the intended knowledge or skills are actually applied” (*The Standards for Educational and Psychological Measurement* (AERA, APA & NCME 1999, p.137). A close proximity between language assessment tasks and what is being assessed is particularly appealing to both language testers and teachers in the era of communicative language teaching and assessment (Yu, in press 2). On the one hand, the performance

assessment tasks as well as the evaluation criteria, if designed appropriately to mirror the construct of interest, can be readily adapted and used in language teaching and learning. On the other hand, performance assessment tasks should also be “curriculum-embedded”, in order to create a system “in which teaching, assessing, record keeping, criticizing, evaluating, exhibiting, and reflecting all serve to enable and enhance learning”, and a system that treats teaching, learning, and assessment as “continuously interacting components, utilizing instructional materials to provide opportunities for assessment and assessment procedures as instruments for instruction” (Gordon & Bonilla-Bowman, 1996, p. 36).

The integration of performance assessment tasks, speaking and writing in particular, with pedagogical tasks, should be bi-lateral. In other words, it is equally desirable to integrate teacher-initiated, classroom-embedded pedagogical performance tasks with standardized large-scale, high-stakes assessments. For example, the School-based Assessment (SBA) in Hong Kong secondary education is such an attempt to use teacher-initiated performance assessments to supplement high-stakes mandated assessments (see Vol. 43, Issue 3, 2009 of *TESOL Quarterly*, a special issue on SBA).

3. Improving the assessment literacy of teachers and the language classroom teaching experience of test designers as part of their respective professional development

In addition to (1) designing more authentic and engaging assessment tasks that reflect current thinking about communicative language teaching and learning, and (2) integrating performance tasks in assessment and instruction, the willingness of both communities to stand in each other's shoes, rather than blaming tests as the source of all problems, can improve the relevance of standardized tests to language teaching and learning.

Teachers' lack of assessment literacy and lack of opportunities to improve their assessment literacy have regularly featured in academic commentary (e.g., Jin, 2010). For language teachers, their professional development should include efforts to better understand the issues and challenges in assessment, from the basic technical testing expertise, knowledge about language acquisition and measurement, to key principles of fairness, ethics, impact and professionalism in large-scale language assessment (Davies, 2008). For language test designers, a better understanding and experience of teaching and learning the target language and the use of formative in-flight assessments (Rea-Dickins 2008) in classroom contexts should be an integral part of their professional development efforts.

I would argue that test providers and assessment professionals have the responsibility to share with teachers the knowledge, skills and understanding that underpin good quality assessment for the benefit of all (see Taylor, 2009). Similarly, it is imperative and beneficial for assessment professionals including assessment task designers to gain more experience in teaching and learning the target language in classroom contexts so as to better understand how assessments, especially, formative classroom-based assessments, are conducted by teachers. Teachers' expertise or literacy in classroom-based assessment can inform and improve the design of performance assessment tasks in large-scale, high-stakes language tests, which in turn can be used readily as “instruments for instruction” (Gordon & Bonilla-Bowman, 1996, p. 36).

4. Concluding remarks

In this paper, I've argued in favour of dialogue between assessment and teaching professionals in our joint efforts to create a system in which standardized performance assessment tasks can be readily used for instructional purposes and in which there does not have to be a clear-cut boundary between teaching and assessment. Assessment tasks in standardized tests must be designed in such a way that they reflect current understanding of communicative language teaching and learning, and such understanding must be equally evidenced in classroom teaching and learning activities. In order to achieve a happy reconciliation between standardized tests, teaching and learning, I've argued for the joint efforts of testing and teaching professionals to be better equipped with each other's professional know-how. If the current atmosphere of blame and recrimination continues, language education will be weakened. But by working together, the two professional communities can promote and embed the relevance of standardized tests within classroom teaching and learning and we will all gain from it.

References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280–297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38–54.
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- Gordon, E. W., & Bonilla-Bowman, C. (1996). Can performance-based assessments contribute to the achievement of educational equity? In J. B. Baron & D. P. Wolf (Eds.), *Performance-based student assessment: Challenges and possibilities. Ninety-fifth yearbook of the National Society for the Study of Education, part i* (pp. 32–51). Chicago: University of Chicago Press.
- Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11(2), 113–134.
- Green, A. (2007). *IELTS Washback in Context*. Cambridge: Cambridge University Press.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295–303.
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584.
- Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Paris: UNESCO: International Institute for Educational Planning.
- McNamara, T. F. (1998). Policy and social considerations in language assessment. *Annual Review of Applied Linguistics*, 18, 304–319.
- McNamara, T. F. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333–349.
- Rea-Dickins, P. M. (1997). So why do we need relationships with stakeholders in language testing. A view from the UK. *Language Testing*, 14(3), 304–314.
- Rea-Dickins, P. M. (2008). Classroom-based language assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.): Vol. 7. *Language testing and assessment* (pp. 257–271). New York: Springer Science+Business Media.
- Rea-Dickins, P. M., Kiely, R., & Yu, G. (2007). Student identity, learning and progression: The affective and academic impact of IELTS on ‘successful’ candidates. In P. McGovern & S. Walsh (Eds.), *IELTS research reports volume 7* (pp. 59–136). Canberra: IELTS Australia and British Council.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298–317.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21–36.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The sri lankan impact study. *Language Testing*, 10(1), 41–69.
- Yu, G. (in press) Preparing for TOEFL iBT speaking tests – a journey of Chinese test takers. *Assessment in Education*. Special Issue.
- Yu, G. (in press 2). Performance assessment. In Kunnan, A. (ed.) *Companion to Language Assessment*. Blackwell-Wiley.